

**Remarks**

Claims 28-30 have been newly added. Reconsideration and allowance of the pending claims are respectfully requested.

**Drawings objection**

The Office Action objected to the drawings because they are handwritten. Applicant has submitted herewith corrected drawing sheets of Figs. 1-6 in compliance with 37 CFR 1.121 (d). Withdrawal of the present objection of drawings is earnestly requested.

**Claims Rejections Under 35 U.S.C. 103(a)**

**Claims 1, 3, 4, 6-9, 11-17, 19, 20 and 22-24**

The Office Action rejected claims 1, 3, 4, 6-9, 11-17, 19, 20 and 22-24 under 35 U.S.C. 103 as being unpatentable over Castellanos et al. (US PAT PUB 2003/0014448A1) in view of Razin et al. (US PAT. 6,098,034). Applicant respectfully requests the rejection of claims 1, 3, 4, 6-9, 11-17, 19, 20 and 22-24 be withdrawn for the following reasons.

As discussed in M.P.E.P 2143.03, to establish prima facie obviousness of a claimed invention, all the claim limitations must be taught or suggested by the prior art. In re Royka, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). "All words in a claim must be considered in judging the patentability of that claim against the prior art." In re Wilson, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970).

Each of claims 1, 9 and 17 comprise **filtering the set of frequently occurring phrases to determine** a set of frequently occurring and **unrecognized phrases** as entity name and jargon term candidates, which is neither taught nor suggested by a combination of Castellanos and Razin.

Instead, Castellanos teaches a method of normalizing dirty text in documents by creating a thesaurus made of standard terms and their variations and passing the documents through the thesaurus for normalization. The thesaurus is setup by the following steps: generating a reference word list and non-reference word list with non-stop words; generating an approximate duplicates list consisting of non-reference words sufficiently close to reference words; manually filtering the approximate duplicates list to setup the thesaurus (See, Figs. 2-3 and corresponding paragraphs [0030]-[0041]).

The Final Office Action appears to rely on paragraphs 0031-0032 on the teaching of filtering the set of frequently occurring phrases to determine a set of frequently occurring and unrecognized phrases. Applicant respectfully submits that paragraphs 0031-0032 teaches a method of generating a reference word list and non-reference word list, comprising: first, placing special words (words with special characters embedded within) and correctly spelled words in the reference word list, while placing incorrectly spelled words in the non-reference words list; second, setting frequency thresholds to keep words in the reference words list and non-reference list respectively, wherein the threshold for the reference list is higher than that for the non-reference list. Therefore, it can be concluded that Castellanos teaches **filtering**

**the set of special or correctly spelled words** with a frequency threshold **to leave a set of frequently occurring** special or correctly spelled **words on the list** (please further refer to claim 4 of Castellanos), but not **filtering the set of frequently occurring phrases to determine** a set of frequently occurring and **unrecognized phrases** as entity name and jargon term candidates, as claimed by each of claims 1, 9, 17 of the present application.

As pointed in the response mailed on September 6, 2005, Razin teaches a method of standardizing phrases in a document, which establishes a master phrase list by filtering the document to determine master phrases meeting predetermined criteria on phrase length and phrase occurring frequency and standardizes phrases in the document by finding and editing phrases approximate to the master phrases. However, Razin does not teach **filtering the set of frequently occurring phrases to determine** a set of frequently occurring and **unrecognized phrases** as entity name and jargon term candidates, as claimed by each of claims 1, 9 and 17 of the present application.

Since neither Castellanos nor Razin teaches or suggests **filtering the set of frequently occurring phrases to determine** a set of frequently occurring and **unrecognized phrases** as entity name and jargon term candidates, the combination thereof does not meet the requirements of an obvious rejection. Applicant respectfully request the present rejection of claims 1, 9 and 17 be withdrawn.

Each of claims 3-4, 6-8, 11-16, 19-20 and 22-27 includes one of claims 1, 9 and 17 as a base claim. Accordingly, each of claims 3-4, 6-8, 11-16, 19-20 and 22-27 is at least allowable for the reasons noted above.

**Claims 2, 10 and 18**

The Office Action rejects claims 2, 10 and 18 under 35. U.S.C. 103 as being unpatentable over Castellanos in view of Razin et al. and taken in further view of Applicant's admitted prior art. Each of claims 2, 10 and 18 include one of claims 1, 10 and 17 as a base claim and are therefore allowable for at least the reasons stated above. Applicant respectfully requests the present rejection of claims 2, 10 and 18 be withdrawn.

**Claims 5, 13 and 21**

The Office Action rejects claims 5, 13 and 21 under 35. U.S.C. 103 as being unpatentable over Castellanos in view of Razin et al. and taken in view of Chien ("PAT-Tree-Based Keyword Extraction for Chinese Information Retrieval"). Each of claims 5, 13 and 21 include one of claims 1, 10 and 17 as a base claim and are therefore allowable for at least the reasons stated above. Applicant respectfully requests the present rejection of claims 5, 13 and 21 be withdrawn.

**Newly Added Claims**

The Applicants have added claims 28-30 which contain limitations not disclosed, taught, or suggested by the cited art. In particular, each of claims 28-30 recites the filtering in each of claims 1, 9 and 17 further comprises excluding an

embedded phrase from the set of frequently occurring phrases, wherein the embedded phrase is embedded by an embedding phrase that occurs at a similar frequency with the embedded phrase.

In order to expedite prosecution, Applicant respectfully point out although Castellanos teaches a method of normalizing dirty text in a document by using a thesaurus established based upon a reference list and non-reference list, Castellanos is silent on embedded phrases, and thus says nothing on how to deal with embedded phrases depending on its occurring frequency. Further, although Razin teaches a method of processing nested phrases, the method of Razin is different from that claimed by each of claims 28-30. Namely, Razin teaches discarding the nested phrase if not occurring outside the nesting phrase a certain number of times, but each of claims 28-30 teaches excluding an embedded phrase if the occurring frequency of the embedded phrase is similar with its embedding phrase.

Therefore, each of claims 28-30 which recites the filtering in each of claims 1, 9 and 17 further comprises excluding an embedded phrase from the set of frequently occurring phrases, wherein the embedded phrase is embedded by an embedding phrase that occurs at a similar frequency with the embedded phrase, is neither taught nor suggested by a combination of Castellanos and Razin.

**Conclusion**

The foregoing is submitted as a full and complete response to the Official Action. Applicant submits that the application is in condition for allowance. Reconsideration is requested, and allowance of the pending claims is earnestly solicited.

Should it be determined that an additional fee is due under 37 CFR §§1.16 or 1.17, or any excess fee has been received, please charge that fee or credit the amount of overcharge to deposit account #02-2666. If the Examiner believes that there are any informalities, which can be corrected by an Examiner's amendment, a telephone call to the undersigned at (503) 439-8778 is respectfully solicited.

Respectfully submitted,



Paul A. Mendonsa  
Reg. No. 42,879

**Blakely, Sokoloff, Taylor & Zafman, LLP**  
12400 Wilshire Blvd.  
Seventh Floor  
Los Angeles, CA 90025-1030  
(408) 720-8300